GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: December 27, 2001, 16:55:45 ; Search time 427.16 Seconds
(without alignments)
5579.558 Million cell updates/sec

Title: US-09-830-647-3
Perfect score: 2780
Sequence: 1 aattcggcacgagctctctg.....aaaaaaaaaaaaaactcgag 2780

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2780	100.0	2780	21	AAA53483	Human H37 nucleoti
	2	2768	99.6	6712	21	AAC76691	Human ORFX ORF2246
	3	2648	95.3	2719	21	AAA53484	Human H37 nucleoti
c	4	1194.2	43.0	1915	22	AAI13747	Probe #3680 for ge
c	5	1194.2	43.0	1915	22	AAI03627	Probe #3618 used t
	6	1101.6	39.6	2276	22	AAH17747	Human cDNA sequenc
c	7	965	34.7	1031	22	AAI22968	Probe #12901 for g
c	8	965	34.7	1031	22	AAI08634	Probe #8625 used t
	9	721.6	26.0	903	22	AAH06140	Human cDNA clone (
	10	513.8	18.5	685	22	AAF64883	Novel human polynu
	11	411	14.8	422	22	AAF66758	Novel human polynu
	12	384.4	13.8	403	22	AAF64912	Novel human polynu
	13	294.6	10.6	318	21	AAA43876	Human secreted exp
	14	211.4	7.6	397	22	AAF64832	Novel human polynu
	15	184.2	6.6	468	22	AAI16129	Probe #6062 for ge
	16	184.2	6.6	468	22	AAI38399	Probe #7085 used t
	17	183.2	6.6	201	22	AAI25319	Probe #15252 for g
	18	183.2	6.6	201	22	AAI51493	Probe #20179 used
c	19	182.2	6.6	416	22	AAI37857	Probe #6543 used t
	20	178.2	6.4	449	22	AAI15475	Probe #5408 for ge
	21	178.2	6.4	449	22	AAI36863	Probe #5549 used t
c	22	176.8	6.4	936	22	AAF58252	Oligonucleotide D1
c	23	176.8	6.4	936	22	AAF58254	Oligonucleotide D1
c	24	176.8	6.4	936	22	AAF58257	Oligonucleotide D1
c	25	176.8	6.4	936	22	AAF58259	Oligonucleotide D2
c	26	176.8	6.4	936	22	AAF58262	Oligonucleotide D2
c	27	176.8	6.4	938	22	AAF58255	Oligonucleotide D1
	28	176.2	6.3	936	22	AAF58252	Oligonucleotide D1
	29	176.2	6.3	936	22	AAF58254	Oligonucleotide D1
	30	176.2	6.3	936	22	AAF58257	Oligonucleotide D1
	31	176.2	6.3	936	22	AAF58259	Oligonucleotide D2
	32	176.2	6.3	936	22	AAF58262	Oligonucleotide D2
	33	176.2	6.3	938	22	AAF58255	Oligonucleotide D1
	34	175.8	6.3	180	16	AAT25803	Human gene signatu
c	35	165	5.9	165	22	AAI57140	Probe #25826 used
	36	149	5.4	149	22	AAI24703	Probe #14636 for g
	37	149	5.4	149	22	AAI50060	Probe #18746 used
c	38	124	4.5	195	21	AAC04338	Human secreted pro
	39	99.4	3.6	377	22	AAF65378	Novel human polynu
c	40	72	2.6	244	22	AAF58238	Oligonucleotide D1
	41	70.4	2.5	244	22	AAF58238	Oligonucleotide D1
c	42	54.8	2.0	12001	16	AAQ76213	HSV L/ST region.
c	43	53	1.9	114793	22	AAD08215	Human genome from
c	44	52.4	1.9	3198	20	AAX02974	Human IL-1ra BAC c
c	45	52	1.9	2188	20	AAZ77506	Human ovarian tumo

ALIGNMENTS

RESULT 1
AAA53483
ID AAA53483 standard; cDNA; 2780 BP.
XX
AC AAA53483;
XX
DT 04-OCT-2000 (first entry)
XX
DE Human H37 nucleotide sequence #1.
XX
KW H37; human; Cdc7 regulatory subunit; cytostatic; proliferative; cancer;
KW anti-proliferative; replication regulator; stem cell; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 518..2542
FT /*tag= a
FT /product= "H37"
XX
PN WO200026250-A1.
XX
PD 11-MAY-2000.
XX
PF 01-NOV-1999; 99WO-JP06076.
XX
PR 30-OCT-1998; 98JP-0311408.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA (ARAI/) ARAI K.
PA (MASA/) MASAI H.
XX
PI Arai K, Masai H;
XX

DR WPI; 2000-365580/31.
DR P-PSDB; AAB03758.
XX
PT Human H37 proteins with a Cdc7 activity regulatory subunit, for
PT controlling cell replication and cell proliferation, useful in treating
PT cancers and diseases due to abnormal production of stem cells -
XX
PS Claim 6; Page 47-49; 55pp; Japanese.
XX
CC The present sequence represents a human H37 encoding nucleotide sequence.
CC H37 is a protein with a Cdc7 activity regulatory subunit. The invention
CC relates to two H37 protein and nucleotide sequences. H37 proteins exhibit
CC cytostatic, proliferative, anti-proliferative, and cell replication
CC regulatory activities. The proteins, encoded genes and DNA fragments are
CC useful in treating cancers and other diseases resulting from abnormal
CC production of stem cells. Antibodies directed against one of the H37
CC proteins can be used to inhibit cell proliferation.
XX
SQ Sequence 2780 BP; 949 A; 534 C; 613 G; 684 T; 0 other;

Query Match 100.0%; Score 2780; DB 21; Length 2780;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 aattcggcacgagctctctgaggctgcgccaagacctgaagcggcggaccgagagcccgg 60
Db       1 aattcggcacgagctctctgaggctgcgccaagacctgaagcggcggaccgagagcccgg 60

Qy      61 gtctgagactgagagagcaacggaatggaggcggggtagaggcggaaacacaacctgcag 120
Db      61 gtctgagactgagagagcaacggaatggaggcggggtagaggcggaaacacaacctgcag 120

Qy     121 ggccagagcgaggcgcgagaaggacggcggcgtgaggggcggggcgcgcagcgcgagaa 180
Db     121 ggccagagcgaggcgcgagaaggacggcggcgtgaggggcggggcgcgcagcgcgagaa 180

Qy     181 ggcaggcacgaggggcgagcgcgaggcggggcacggcgcgtggcgtgagacggggcgggg 240
Db     181 ggcaggcacgaggggcgagcgcgaggcggggcacggcgcgtggcgtgagacggggcgggg 240

Qy     241 cgcgcgtatcggcgccgcggccgcgtgacgcgttttcaaatcttcaaccgccgcagccca 300
Db     241 cgcgcgtatcggcgccgcggccgcgtgacgcgttttcaaatcttcaaccgccgcagccca 300

Qy     301 ctcgtttgtgctttgcgccttcctcctccgcgccttggagccggatccggccccggaaac 360
Db     301 ctcgtttgtgctttgcgccttcctcctccgcgccttggagccggatccggccccggaaac 360

Qy     361 ccgacctgcagacgcggtacctctactgcgtagaggccgtagctggcggaaggagagagg 420
Db     361 ccgacctgcagacgcggtacctctactgcgtagaggccgtagctggcggaaggagagagg 420

Qy     421 cggccgtcctgtcaacaggccgggggaagccgtgctttcgcggctgcccggtgcgacact 480
Db     421 cggccgtcctgtcaacaggccgggggaagccgtgctttcgcggctgcccggtgcgacact 480

Qy     481 ttctccggacccagcatgtaggtgccgggcgactgccatgaactccggagccatgaggat 540
Db     481 ttctccggacccagcatgtaggtgccgggcgactgccatgaactccggagccatgaggat 540

Qy     541 ccacagtaaaggacatttccagggtggaatccaagtcaaaaatgaaaaaaacagaccatc 600
Db     541 ccacagtaaaggacatttccagggtggaatccaagtcaaaaatgaaaaaaacagaccatc 600

Qy     601 tctgaaatctctgaaaactgataacaggccagaaaaatccaaatgtaagccactttgggg 660
Db     601 tctgaaatctctgaaaactgataacaggccagaaaaatccaaatgtaagccactttgggg 660

Qy     661 aaaagtattttaccttgacttaccttctgtcaccatatctgaaaaacttcaaaaggacat 720
Db     661 aaaagtattttaccttgacttaccttctgtcaccatatctgaaaaacttcaaaaggacat 720

Qy     721 taaggatctgggagggcgagttgaagaatttctcagcaaagatatcagttatcttatttc 780
Db     721 taaggatctgggagggcgagttgaagaatttctcagcaaagatatcagttatcttatttc 780

Qy     781 aaataagaaggaagctaaatttgcacaaaccttgggtcgaatttctcctgtaccaagtcc 840
Db     781 aaataagaaggaagctaaatttgcacaaaccttgggtcgaatttctcctgtaccaagtcc 840

Qy     841 agaatctgcatatactgcagaaaccacttcacctcatcccagccatgatggaagttcatt 900
Db     841 agaatctgcatatactgcagaaaccacttcacctcatcccagccatgatggaagttcatt 900

Qy     901 taagtcaccagacacagtgtgtttaagcagaggaaaattattagttgaaaaagctatcaa 960
Db     901 taagtcaccagacacagtgtgtttaagcagaggaaaattattagttgaaaaagctatcaa 960

Qy     961 ggaccatgattttattccttcaaatagtatattatcaaatgccttgtcatggggagtaaa 1020
Db     961 ggaccatgattttattccttcaaatagtatattatcaaatgccttgtcatggggagtaaa 1020

Qy    1021 aattcttcatattgatgacattagatactacattgaacaaaagaaaaaagagttgtattt 1080
Db    1021 aattcttcatattgatgacattagatactacattgaacaaaagaaaaaagagttgtattt 1080

Qy    1081 actcaagaaatcaagtacttcagtaagagatgggggcaaaagagttggtagtggtgcaca 1140
Db    1081 actcaagaaatcaagtacttcagtaagagatgggggcaaaagagttggtagtggtgcaca 1140

Qy    1141 aaaaacaagaacaggaagactcaaaaagccttttgtaaaggtggaagatatgagccaact 1200
Db    1141 aaaaacaagaacaggaagactcaaaaagccttttgtaaaggtggaagatatgagccaact 1200

Qy    1201 ttataggccattttatcttcagctgaccaatatgccttttataaattattctattcagaa 1260
Db    1201 ttataggccattttatcttcagctgaccaatatgccttttataaattattctattcagaa 1260

Qy    1261 gccctgcagtccatttgatgtagacaagccatctagtatgcaaaagcaaactcaggttaa 1320
Db    1261 gccctgcagtccatttgatgtagacaagccatctagtatgcaaaagcaaactcaggttaa 1320

Qy    1321 actaagaatccaaacagatggcgataagtatggtggaacctcaattcaactccagttgaa 1380
Db    1321 actaagaatccaaacagatggcgataagtatggtggaacctcaattcaactccagttgaa 1380

Qy    1381 agagaagaagaaaaaaggatattgtgaatgttgcttgcagaaatatgaagatctagaaac 1440
Db    1381 agagaagaagaaaaaaggatattgtgaatgttgcttgcagaaatatgaagatctagaaac 1440

Qy    1441 tcaccttctaagtgagcaacacagaaactttgcacagagtaaccagtatcaagttgttga 1500
Db    1441 tcaccttctaagtgagcaacacagaaactttgcacagagtaaccagtatcaagttgttga 1500

Qy    1501 tgatattgtatctaagttagtttttgactttgtggaatatgaaaaggacacacctaaaaa 1560
Db    1501 tgatattgtatctaagttagtttttgactttgtggaatatgaaaaggacacacctaaaaa 1560

Qy    1561 gaaaagaataaaatacagtgttggatccctttctcctgtttctgcaagtgtcctgaaaaa 1620
Db    1561 gaaaagaataaaatacagtgttggatccctttctcctgtttctgcaagtgtcctgaaaaa 1620

Qy    1621 gactgaacaaaaggaaaaagtggaattgcaacatatttctcagaaagattgccaggaaga 1680
Db    1621 gactgaacaaaaggaaaaagtggaattgcaacatatttctcagaaagattgccaggaaga 1680

Qy    1681 tgatacaacagtgaaggagcagaatttcctgtataaagagacccaggaaactgaaaaaaa 1740
Db    1681 tgatacaacagtgaaggagcagaatttcctgtataaagagacccaggaaactgaaaaaaa 1740

Qy    1741 gctcctgtttatttcagagcccatccccccaccccttcaaatgaattgagagggcttaatga 1800
Db    1741 gctcctgtttatttcagagcccatccccccaccccttcaaatgaattgagagggcttaatga 1800

Qy    1801 gaaaatgagtaataaatgttccatgttaagtacagctgaagatgacataagacagaattt 1860
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 gaaaatgagtaataaatgttccatgttaagtacagctgaagatgacataagacagaatttt 1860

Qy    1861 tacacagctacctctacataaaaacaaacaggaatgcattcttgacatttccgaacacac 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 tacacagctacctctacataaaaacaaacaggaatgcattcttgacatttccgaacacac 1920

Qy    1921 attaagtgaaaatgacttagaagaactaagggtagatcactataaatgtaacatacaggc 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 attaagtgaaaatgacttagaagaactaagggtagatcactataaatgtaacatacaggc 1980

Qy    1981 atctgtacatgtttctgatttcagtacagataatagtggatctcaaccaaaaacagaagtc 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 atctgtacatgtttctgatttcagtacagataatagtggatctcaaccaaaaacagaagtc 2040

Qy    2041 agatactgtgctttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttac 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 agatactgtgctttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttac 2100

Qy    2101 tcatgattctggtctgataacaataaacagttcacaagagcacctaactgttcaggcaaa 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 tcatgattctggtctgataacaataaacagttcacaagagcacctaactgttcaggcaaa 2160

Qy    2161 ggctccattccatactcctcctgaggaacccaatgaatgtgacttcaagaatatggatag 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 ggctccattccatactcctcctgaggaacccaatgaatgtgacttcaagaatatggatag 2220

Qy    2221 tttaccttctggtaaaatacatcgaaaagtgaaaataatattaggacgaaatagaaaaga 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 tttaccttctggtaaaatacatcgaaaagtgaaaataatattaggacgaaatagaaaaga 2280

Qy    2281 aaatctggaaccaaatgctgaatttgataaaagaactgaatttattacacaagaagaaaa 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 aaatctggaaccaaatgctgaatttgataaaagaactgaatttattacacaagaagaaaa 2340

Qy    2341 cagaatttgtagttcaccggtacagtctttactagacttgtttcagactagtgaagagaa 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 cagaatttgtagttcaccggtacagtctttactagacttgtttcagactagtgaagagaa 2400

Qy    2401 atcagaatttttgggtttcacaagctacacagaaaagagtggtatatgcaatgttttaga 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 atcagaatttttgggtttcacaagctacacagaaaagagtggtatatgcaatgttttaga 2460

Qy    2461 tatttgggaagaggaaaattcagataatctgttaacagcgtttttctcgtccccttcaac 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 tatttgggaagaggaaaattcagataatctgttaacagcgtttttctcgtccccttcaac 2520

Qy    2521 ttctacatttactggctttagaatttaaaaaatgcatacttttcagaagtgataaggat 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 ttctacatttactggctttagaatttaaaaaatgcatacttttcagaagtgataaggat 2580

Qy    2581 catattcttgaaatttttataaatatgtatggaaattcttaggatttttttaccagcttt 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 catattcttgaaatttttataaatatgtatggaaattcttaggatttttttaccagcttt 2640

Qy    2641 gtttacagacccaaatgtaaatattaaaaataaatatttgcaatttttctacagaattgaa 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 gtttacagacccaaatgtaaatattaaaaataaatatttgcaatttttctacagaattgaa 2700

Qy    2701 tacctgttaaagaaaaattacagaataaacttgtgactggtcttgttttacattaaaaaa 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 tacctgttaaagaaaaattacagaataaacttgtgactggtcttgttttacattaaaaaa 2760

Qy    2761 aaaaaaaaaaaaaactcgag 2780
           ||||||||||||||||||||
Db    2761 aaaaaaaaaaaaaactcgag 2780


RESULT   2
AAC76691
ID   AAC76691 standard; cDNA; 6712 BP.
XX
AC   AAC76691;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human ORFX ORF2246 polynucleotide sequence SEQ ID NO:4491.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive; ss.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   P-PSDB; AAB42482.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 5; Page 3671-3675; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence 6712 BP; 2319 A; 1206 C; 1244 G; 1942 T; 1 other;


  Query Match             99.6%;  Score 2768;  DB 21;  Length 6712;
  Best Local Similarity   99.8%;  Pred. No. 0;
```

Matches 2771; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

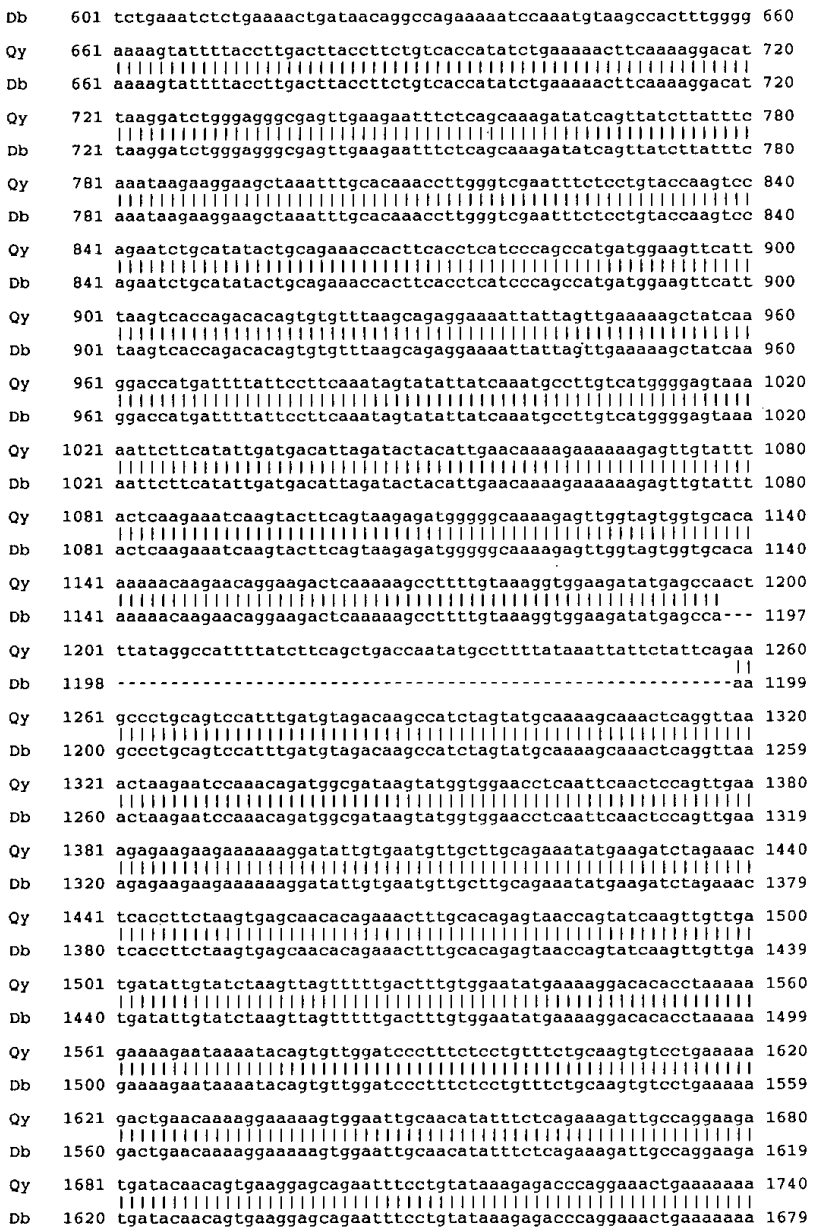
```
Qy       5 cggcacgagctctctgaggctgcgccaagacctgaagcggcggaccgagagcccgggtct 64
Db    3937 ccgcgccaactctctgaggctgcgccaagacctgaagcggcggaccgagagcccgggtct 3996

Qy      65 gagactgagagagcaacggaatggaggcggggtagaggcggaaacacaacctgcagggcc 124
Db    3997 gagactgagagagcaacggaatggaggcggggtagaggcggaaacacaacctgcagggcc 4056

Qy     125 agagcgaggcgcgagaaggacggcggcgtgaggggggcgggcgcgcagcgcgagaaggca 184
Db    4057 agagcgaggcgcgagaaggacggcggcgtgaggggggcgggcgcgcagcgcgagaaggca 4116

Qy     185 ggcacgaggggcgagcgcgaggcggggcacggcgcgtggcgtgagacggggcggggcgcg 244
Db    4117 ggcacgaggggcgagcgcgaggcggggcacggcgcgtggcgtgagacggggcggggcgcg 4176

Qy     245 cgtatcggcgccgcggccgcgtgacgcgttttcaaatcttcaaccgccgcagcccactcg 304
Db    4177 cgtatcggcgccgcggccgcgtgacgcgttttcaaatcttcaaccgccgcagcccactcg 4236

Qy     305 tttgtgctttgcgccttcctcctccgcgccttggagccggatccggccccggaaacccga 364
Db    4237 tttgtgctttgcgccttcctcctccgcgccttggagccggatccggccccggaaacccga 4296

Qy     365 cctgcagacgcggtacctctactgcgtagaggccgtagctggcggaaggagagaggcggc 424
Db    4297 cccgcagacgcggtacctctactgcgtagaggccgtagctggcggaaggagagaggcggc 4356

Qy     425 cgtcctgtcaacaggccgggggaagccgtgctttcgcggctgcccggtgcgacactttct 484
Db    4357 cgtcctgtcaacaggccgggggaagccgtgctttcgcggctgcccggtgcgacactttct 4416

Qy     485 ccggacccagcatgtaggtgccgggcgactgccatgaactccggagccatgaggatccac 544
Db    4417 ccggacccagcatgtaggtgccgggcgactgccatgaactccggagccatgaggatccac 4476

Qy     545 agtaaaggacatttccagggtggaatccaagtcaaaaatgaaaaaaacagaccatctctg 604
Db    4477 agtaaaggacatttccagggtggaatccaagtcaaaaatgaaaaaaacagaccatctctg 4536

Qy     605 aaatctctgaaaactgataacaggccagaaaaatccaaatgtaagccactttggggaaaa 664
Db    4537 aaatctctgaaaactgataacaggccagaaaaatccaaatgtaagccactttggggaaaa 4596

Qy     665 gtattttaccttgacttaccttctgtcaccatatctgaaaaacttcaaaaggacattaag 724
Db    4597 gtattttaccttgacttaccttctgtcaccatatctgaaaaacttcaaaaggacattaag 4656

Qy     725 gatctgggagggcgagttgaagaatttctcagcaaagatatcagttatcttatttcaaat 784
Db    4657 gatctgggagggcgagttgaagaatttctcagcaaagatatcagttatcttatttcaaat 4716

Qy     785 aagaaggaagctaaatttgcacaaaccttgggtcgaatttctcctgtaccaagtccagaa 844
Db    4717 aagaaggaagctaaatttgcacaaaccttgggtcgaatttctcctgtaccaagtccagaa 4776

Qy     845 tctgcatatactgcagaaaccacttcacctcatcccagccatgatggaagttcatttaag 904
Db    4777 tctgcatatactgcagaaaccacttcacctcatcccagccatgatggaagttcatttaag 4836

Qy     905 tcaccagacacagtgtgtttaagcagaggaaaattattagttgaaaaagctatcaaggac 964
Db    4837 tcaccagacacagtgtgtttaagcagaggaaaattattagttgaaaaagctatcaaggac 4896

Qy     965 catgattttattccttcaaatagtatattatcaaatgccttgtcatggggagtaaaaatt 1024
Db    4897 catgattttattccttcaaatagtatattatcaaatgccttgtcatggggagtaaaaatt 4956

Qy    1025 cttcatattgatgacattagatactacattgaacaaaagaaaaaagagttgtatttactc 1084
Db    4957 cttcatattgatgacattagatactacattgaacaaaagaaaaaagagttgtatttactc 5016

Qy    1085 aagaaatcaagtacttcagtaagagatgggggcaaaagagttggtagtggtgcacaaaaa 1144
Db    5017 aagaaatcaagtacttcagtaagagatgggggcaaaagagttggtagtggtgcacaaaaa 5076

Qy    1145 acaagaacaggaagactcaaaaagccttttgtaaaggtggaagatatgagccaactttat 1204
Db    5077 acaagaacaggaagactcaaaaagccttttgtaaaggtggaagatatgagccaactttat 5136

Qy    1205 aggccattttatcttcagctgaccaatatgccttttataaattattctattcagaagccc 1264
Db    5137 aggccattttatcttcagctgaccaatatgccttttataaattattctattcagaagccc 5196

Qy    1265 tgcagtccatttgatgtagacaagccatctagtatgcaaaagcaaactcaggttaaacta 1324
Db    5197 tgcagtccatttgatgtagacaagccatctagtatgcaaaagcaaactcaggttaaacta 5256

Qy    1325 agaatccaaacagatggcgataagtatggtggaacctcaattcaactccagttgaaagag 1384
Db    5257 agaatccaaacagatggcgataagtatggtggaacctcaattcaactccagttgaaagag 5316

Qy    1385 aagaagaaaaaaggatattgtgaatgttgcttgcagaaatatgaagatctagaaactcac 1444
Db    5317 aagaagaaaaaaggatattgtgaatgttgcttgcagaaatatgaagatctagaaactcac 5376

Qy    1445 cttctaagtgagcaacacagaaactttgcacagagtaaccagtatcaagttgttgatgat 1504
Db    5377 cttctaagtgagcaacacagaaactttgcacagagtaaccagtatcaagttgttgatgat 5436

Qy    1505 attgtatctaagttagtttttgactttgtggaatatgaaaaggacacacctaaaaagaaa 1564
Db    5437 attgtatctaagttagtttttgactttgtggaatatgaaaaggacacacctaaaaagaaa 5496

Qy    1565 agaataaaatacagtgttggatccctttctcctgtttctgcaagtgtcctgaaaaagact 1624
Db    5497 agaataaaatacagtgttggatccctttctcctgtttctgcaagtgtcctgaaaaagact 5556

Qy    1625 gaacaaaaggaaaaagtggaattgcaacatatttctcagaaagattgccaggaagatgat 1684
Db    5557 gaacaaaaggaaaaagtggaattgcaacatatttctcagaaagattgccaggaagatgat 5616

Qy    1685 acaacagtgaaggagcagaatttcctgtataaagagacccaggaaactgaaaaaaagctc 1744
Db    5617 acaacagtgaaggagcagaatttcctgtataaagagacccaggaaactgaaaaaaagctc 5676

Qy    1745 ctgtttatttcagagcccatcccccaccctcaaatgaattgagagggcttaatgagaaa 1804
Db    5677 ctgtttatttcagagcccatcccccaccctcaaatgaattgagagggcttaatgagaaa 5736

Qy    1805 atgagtaataaatgttccatgttaagtacagctgaagatgacataagacagaatttaca 1864
Db    5737 atgagtaataaatgttccatgttaagtacagctgaagatgacataagacagaatttaca 5796

Qy    1865 cagctacctctacataaaaacaaacaggaatgcattcttgacatttccgaacacacatta 1924
Db    5797 cagctacctctacataaaaacaaacaggaatgcattcttgacatttccgaacacacatta 5856

Qy    1925 agtgaaaatgacttagaagaactaagggtagatcactataaatgtaacatacaggcatct 1984
Db    5857 agtgaaaatgacttagaagaactaagggtagatcactataaatgtaacatacaggcatct 5916

Qy    1985 gtacatgtttctgatttcagtacagataatagtggatctcaaccaaaacagaagtcagat 2044
Db    5917 gtacatgtttctgatttcagtacagataatagtggatctcaaccaaaacagaagtcagat 5976

Qy    2045 actgtgctttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttactcat 2104
Db    5977 actgtgctttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttactcat 6036

Qy    2105 gattctggtctgataacaataaacagttcacaagagcacctaactgttcaggcaaaggct 2164
Db    6037 gattctggtctgataacaataaacagttcacaagagcacctaactgttcaggcaaaggct 6096
```

```
Qy   2165 ccattccatactcctcctgaggaacccaatgaatgtgacttcaagaatatggatagttta 2224
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6097 ccattccatactcctcctgaggaacccaatgaatgtgacttcaagaatatggatagttta 6156

Qy   2225 ccttctggtaaaatacatcgaaaagtgaaaataatattaggacgaaatagaaaagaaaat 2284
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6157 ccttctggtaaaatacatcgaaaagtgaaaataatattaggacgaaatagaaaagaaaat 6216

Qy   2285 ctggaaccaaatgctgaatttgataaaagaactgaatttattacacaagaagaaaacaga 2344
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6217 ctggaaccaaatgctgaatttgataaaagaactgaatttattacacaagaagaaaacaga 6276

Qy   2345 atttgtagttcaccggtacagtctttactagacttgtttcagactagtgaagagaaatca 2404
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6277 atttgtagttcaccggtacagtctttactagacttgtttcagactagtgaagagaaatca 6336

Qy   2405 gaatttttgggtttcacaagctacacagaaaagagtggtatatgcaatgttttagatatt 2464
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6337 gaatttttgggtttcacaagctacacagaaaagagtggtatatgcaatgttttagatatt 6396

Qy   2465 tgggaagaggaaaattcagataatctgttaacagcgtttttctcgtccccttcaacttct 2524
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6397 tgggaagaggaaaattcagataatctgttaacagcgtttttctcgtccccttcaacttct 6456

Qy   2525 acatttactggcttttagaatttaaaaaatgcatacttttcagaagtgataaggatcata 2584
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6457 acatttactggcttttagaatttaaaaaatgcatacttttcagaagtgataaggatcata 6516

Qy   2585 ttcttgaaatttttataaatatgtatggaaattcttaggatttttttaccagctttgttt 2644
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6517 ttcttgaaatttttataaatatgtatggaaattcttaggatttttttaccagctttgttt 6576

Qy   2645 acagacccaaatgtaaatattaaaaaataaatatttgcaattttctacagaattgaatacc 2704
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6577 acagacccaaatgtaaatattaaaaaataaatatttgcaattttctacagaattgaatacc 6636

Qy   2705 tgttaaagaaaaattacagaataaacttgtgactggtcttgttttacattaaaaaaaaaa 2764
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6637 tgttaaagaaaaattacagaataaacttgtgactggtcttgttttacattaaaaaaaaaa 6696

Qy   2765 aaaaaaaaaactcgag 2780
          ||||||||||||||||
Db   6697 aaaaaaaaaactcgag 6712
```

RESULT 3
AAA53484
ID AAA53484 standard; cDNA; 2719 BP.
XX
AC AAA53484;
XX
DT 04-OCT-2000 (first entry)
XX
DE Human H37 nucleotide sequence #2.
XX
KW H37; human; Cdc7 regulatory subunit; cytostatic; proliferative; cancer;
KW anti-proliferative; replication regulator; stem cell; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 518..1222
FT /*tag= a
FT /product= "H37"
XX
PN WO200026250-A1.
XX
PD 11-MAY-2000.
XX
PF 01-NOV-1999; 99WO-JP06076.
XX
PR 30-OCT-1998; 98JP-0311408.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA (ARAI/) ARAI K.
PA (MASA/) MASAI H.
XX
PI Arai K, Masai H;
XX
DR WPI; 2000-365580/31.
DR P-PSDB; AAB03759.
XX
PT Human H37 proteins with a Cdc7 activity regulatory subunit, for
PT controlling cell replication and cell proliferation, useful in treating
PT cancers and diseases due to abnormal production of stem cells -
XX
PS Claim 7; Page 49-51; 55pp; Japanese.
XX
CC The present sequence represents a human H37 encoding nucleotide sequence.
CC H37 is a protein with a Cdc7 activity regulatory subunit. The invention
CC relates to two H37 protein and nucleotide sequences. H37 proteins exhibit
CC cytostatic, proliferative, anti-proliferative, and cell replication
CC regulatory activities. The proteins, encoded genes and DNA fragments are
CC useful in treating cancers and other diseases resulting from abnormal
CC production of stem cells. Antibodies directed against one of the H37
CC proteins can be used to inhibit cell proliferation.
XX
SQ Sequence 2719 BP; 932 A; 522 C; 607 G; 658 T; 0 other;

Query Match 95.3%; Score 2648; DB 21; Length 2719;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 0; Indels 61; Gaps 1;

```
Qy      1 aattcggcacgagctctctgaggctgcgccaagacctgaagcggcggaccgagagcccgg 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 aattcggcacgagctctctgaggctgcgccaagacctgaagcggcggaccgagagcccgg 60

Qy     61 gtctgagactgagagagcaacggaatggaggcggggtagaggcggaaacacaacctgcag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gtctgagactgagagagcaacggaatggaggcggggtagaggcggaaacacaacctgcag 120

Qy    121 ggccagagcgaggcgcgagaaggacggcggcgtgaggggcggggcgcgcagcgcgagaa 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ggccagagcgaggcgcgagaaggacggcggcgtgaggggcggggcgcgcagcgcgagaa 180

Qy    181 ggcaggcacgaggggcgagcgcgaggcggggcacggcgcgtggcgtgagacggggcgggg 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ggcaggcacgaggggcgagcgcgaggcggggcacggcgcgtggcgtgagacggggcgggg 240

Qy    241 cgcgcgtatcggcgccgcggccgcgtgacgcgttttcaaatcttcaaccgccgcagccca 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 cgcgcgtatcggcgccgcggccgcgtgacgcgttttcaaatcttcaaccgccgcagccca 300

Qy    301 ctcgtttgtgctttgcgccttcctcctccgcgccttggagccggatccggccccggaaac 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ctcgtttgtgctttgcgccttcctcctccgcgccttggagccggatccggccccggaaac 360

Qy    361 ccgacctgcagacgcggtacctctactgcgtagaggccgtagctggcggaaggagagagg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ccgacctgcagacgcggtacctctactgcgtagaggccgtagctggcggaaggagagagg 420

Qy    421 cggccgtcctgtcaacaggccgggggaagccgtgctttcgcggctgcccggtgcgacact 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cggccgtcctgtcaacaggccgggggaagccgtgctttcgcggctgcccggtgcgacact 480

Qy    481 ttctccggacccagcatgtaggtgccgggcgactgccatgaactccggagccatgaggat 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ttctccggacccagcatgtaggtgccgggcgactgccatgaactccggagccatgaggat 540

Qy    541 ccacagtaaaggacatttccagggtggaatccaagtcaaaaatgaaaaaaacagaccatc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ccacagtaaaggacatttccagggtggaatccaagtcaaaaatgaaaaaaacagaccatc 600

Qy    601 tctgaaatctctgaaaactgataacaggccagaaaaatccaaatgtaagccactttgggg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```


Db 601 tctgaaatctctgaaaactgataacaggccagaaaaatccaaatgtaagccactttgggg 660
Qy 661 aaaagtattttaccttgacttaccttctgtcaccatatctgaaaaacttcaaaaggacat 720
Db 661 aaaagtattttaccttgacttaccttctgtcaccatatctgaaaaacttcaaaaggacat 720
Qy 721 taaggatctgggagggcgagttgaagaatttctcagcaaagatatcagttatcttatttc 780
Db 721 taaggatctgggagggcgagttgaagaatttctcagcaaagatatcagttatcttatttc 780
Qy 781 aaataagaaggaagctaaatttgcacaaaccttgggtcgaatttctcctgtaccaagtcc 840
Db 781 aaataagaaggaagctaaatttgcacaaaccttgggtcgaatttctcctgtaccaagtcc 840
Qy 841 agaatctgcatatactgcagaaaccacttcacctcatcccagccatgatggaagttcatt 900
Db 841 agaatctgcatatactgcagaaaccacttcacctcatcccagccatgatggaagttcatt 900
Qy 901 taagtcaccagacacagtgtgtttaagcagaggaaaattattagttgaaaaagctatcaa 960
Db 901 taagtcaccagacacagtgtgtttaagcagaggaaaattattagttgaaaaagctatcaa 960
Qy 961 ggaccatgattttattccttcaaatagtatattatcaaatgccttgtcatggggagtaaa 1020
Db 961 ggaccatgattttattccttcaaatagtatattatcaaatgccttgtcatggggagtaaa 1020
Qy 1021 aattcttcatattgatgacattagatactacattgaacaaaagaaaaaagagttgtattt 1080
Db 1021 aattcttcatattgatgacattagatactacattgaacaaaagaaaaaagagttgtattt 1080
Qy 1081 actcaagaaatcaagtacttcagtaagagatgggggcaaaagagttggtagtggtgcaca 1140
Db 1081 actcaagaaatcaagtacttcagtaagagatgggggcaaaagagttggtagtggtgcaca 1140
Qy 1141 aaaaacaagaacaggaagactcaaaaagccttttgtaaaggtggaagatatgagccaact 1200
Db 1141 aaaaacaagaacaggaagactcaaaaagccttttgtaaaggtggaagatatgagcca--- 1197
Qy 1201 ttataggccatttatcttcagctgaccaatatgccttttataaattattctattcagaa 1260
Db 1198 ----------aa 1199
Qy 1261 gccctgcagtccatttgatgtagacaagccatctagtatgcaaaagcaaactcaggttaa 1320
Db 1200 gccctgcagtccatttgatgtagacaagccatctagtatgcaaaagcaaactcaggttaa 1259
Qy 1321 actaagaatccaaacagatggcgataagtatggtggaacctcaattcaactccagttgaa 1380
Db 1260 actaagaatccaaacagatggcgataagtatggtggaacctcaattcaactccagttgaa 1319
Qy 1381 agagaagaagaaaaaaggatattgtgaatgttgcttgcagaaatatgaagatctagaaac 1440
Db 1320 agagaagaagaaaaaaggatattgtgaatgttgcttgcagaaatatgaagatctagaaac 1379
Qy 1441 tcaccttctaagtgagcaacacagaaactttgcacagagtaaccagtatcaagttgttga 1500
Db 1380 tcaccttctaagtgagcaacacagaaactttgcacagagtaaccagtatcaagttgttga 1439
Qy 1501 tgatattgtatctaagttagtttttgactttgtggaatatgaaaaggacacacctaaaaa 1560
Db 1440 tgatattgtatctaagttagtttttgactttgtggaatatgaaaaggacacacctaaaaa 1499
Qy 1561 gaaaagaataaaatacagtgttggatccctttctcctgtttctgcaagtgtcctgaaaaa 1620
Db 1500 gaaaagaataaaatacagtgttggatccctttctcctgtttctgcaagtgtcctgaaaaa 1559
Qy 1621 gactgaacaaaaggaaaaagtggaattgcaacatatttctcagaaagattgccaggaaga 1680
Db 1560 gactgaacaaaaggaaaaagtggaattgcaacatatttctcagaaagattgccaggaaga 1619
Qy 1681 tgatacaacagtgaaggagcagaatttcctgtataaagagacccaggaaactgaaaaaaa 1740
Db 1620 tgatacaacagtgaaggagcagaatttcctgtataaagagacccaggaaactgaaaaaaa 1679
Qy 1741 gctcctgtttatttcagagcccatcccccaccccttcaaatgaattgagagggcttaatga 1800
Db 1680 gctcctgtttatttcagagcccatcccccaccccttcaaatgaattgagagggcttaatga 1739
Qy 1801 gaaaatgagtaataaatgttccatgttaagtacagctgaagatgacataagacagaattt 1860
Db 1740 gaaaatgagtaataaatgttccatgttaagtacagctgaagatgacataagacagaattt 1799
Qy 1861 tacacagctacctctacataaaaacaaacaggaatgcattcttgacatttccgaacacac 1920
Db 1800 tacacagctacctctacataaaaacaaacaggaatgcattcttgacatttccgaacacac 1859
Qy 1921 attaagtgaaaatgacttagaagaactaagggtagatcactataaatgtaacatacaggc 1980
Db 1860 attaagtgaaaatgacttagaagaactaagggtagatcactataaatgtaacatacaggc 1919
Qy 1981 atctgtacatgtttctgatttcagtacagataatagtggatctcaaccaaaacagaagtc 2040
Db 1920 atctgtacatgtttctgatttcagtacagataatagtggatctcaaccaaaacagaagtc 1979
Qy 2041 agatactgtgctttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttac 2100
Db 1980 agatactgtgctttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttac 2039
Qy 2101 tcatgattctggtctgataacaataaacagttcacaagagcacctaactgttcaggcaaa 2160
Db 2040 tcatgattctggtctgataacaataaacagttcacaagagcacctaactgttcaggcaaa 2099
Qy 2161 ggctccattccatactcctcctgaggaacccaatgaatgtgacttcaagaatatggatag 2220
Db 2100 ggctccattccatactcctcctgaggaacccaatgaatgtgacttcaagaatatggatag 2159
Qy 2221 tttaccttctggtaaaatacatcgaaaagtgaaaataatattaggacgaaatagaaaaga 2280
Db 2160 tttaccttctggtaaaatacatcgaaaagtgaaaataatattaggacgaaatagaaaaga 2219
Qy 2281 aaatctggaaccaaatgctgaatttgataaaagaactgaatttattacacaagaagaaaa 2340
Db 2220 aaatctggaaccaaatgctgaatttgataaaagaactgaatttattacacaagaagaaaa 2279
Qy 2341 cagaatttgtagttcaccggtacagtctttactagacttgtttcagactagtgaagagaa 2400
Db 2280 cagaatttgtagttcaccggtacagtctttactagacttgtttcagactagtgaagagaa 2339
Qy 2401 atcagaatttttgggtttcacaagctacacagaaaagagtggtatatgcaatgttttaga 2460
Db 2340 atcagaatttttgggtttcacaagctacacagaaaagagtggtatatgcaatgttttaga 2399
Qy 2461 tatttgggaagaggaaaattcagataatctgttaacagcgtttttctcgtccccttcaac 2520
Db 2400 tatttgggaagaggaaaattcagataatctgttaacagcgtttttctcgtccccttcaac 2459
Qy 2521 ttctacatttactggcttttagaatttaaaaaatgcatacttttcagaagtgataaggat 2580
Db 2460 ttctacatttactggcttttagaatttaaaaaatgcatacttttcagaagtgataaggat 2519
Qy 2581 catattcttgaaatttttataaatatgtatggaaattcttaggatttttttaccagcttt 2640
Db 2520 catattcttgaaatttttataaatatgtatggaaattcttaggatttttttaccagcttt 2579
Qy 2641 gtttacagacccaaatgtaaatattaaaaataaatatttgcaattttctacagaattgaa 2700
Db 2580 gtttacagacccaaatgtaaatattaaaaataaatatttgcaattttctacagaattgaa 2639
Qy 2701 tacctgttaaagaaaaattacagaataaacttgtgactggtcttgttttacattaaaaaa 2760
Db 2640 tacctgttaaagaaaaattacagaataaacttgtgactggtcttgttttacattaaaaaa 2699
Qy 2761 aaaaaaaaaaaaaactcgag 2780
Db 2700 aaaaaaaaaaaaaactcgag 2719

```
RESULT   4
AAI13747/c
ID   AAI13747 standard; DNA; 1915 BP.
XX
AC   AAI13747;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #3680 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 3680; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1915 BP; 608 A; 324 C; 280 G; 703 T; 0 other;


  Query Match             43.0%;  Score 1194.2;  DB 22;  Length 1915;
  Best Local Similarity   99.7%;  Pred. No. 5.2e-258;
  Matches 1196;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1563 aaagaataaaatacagtgttggatccctttctcctgtttctgcaagtgtcctgaaaaaga 1622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1283 AAAGAATAAAATACAGTGTTGGATCCCTTTCTCCTGTTTCTGCAAGTGTCCTGAAAAAGA 1224

Qy   1623 ctgaacaaaaggaaaaagtggaattgcaacatatttctcagaaagattgccaggaagatg 1682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1223 CTGAACAAAAGGAAAAAGTGGAATTGCAACATATTTCTCAGAAAGATTGCCAGGAAGATG 1164

Qy   1683 atacaacagtgaaggagcagaatttcctgtataaagagacccaggaaactgaaaaaaagc 1742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1163 ATACAACAGTGAAGGAGCAGAATTTCCTGTATAAAGAGACCCAGGAAACTGAAAAAAAGC 1104

Qy   1743 tcctgtttatttcagagcccatccccccaccttcaaatgaattgagagggcttaatgaga 1802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   1103 TCCTGTTTATTTCAGAGCCCATCCCCCACCCTTCAAATGAATTGAGAGGGCTTAATGAGA 1044

Qy   1803 aaatgagtaataaatgttccatgttaagtacagctgaagatgacataagacagaatttta 1862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1043 AAATGAGTAATAAATGTTCCATGTTAAGTACAGCTGAAGATGACATAAGACAGAATTTTA 984

Qy   1863 cacagctacctctacataaaaacaaacaggaatgcattcttgacatttccgaacacacat 1922
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    983 CACAGCTACCTCTACATAAAAACAAACAGGAATGCATTCTTGACATTTCCGAACACACAT 924

Qy   1923 taagtgaaaatgacttagaagaactaagggtagatcactataaatgtaacatacaggcat 1982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    923 TAAGTGAAAATGACTTAGAAGAACTAAGGGTAGATCACTATAAATGTAACATACAGGCAT 864

Qy   1983 ctgtacatgtttctgatttcagtacagataatagtggatctcaaccaaaaacagaagtcag 2042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    863 CTGTACATGTTTCTGATTTCAGTACAGATAATAGTGGATCTCAACCAAAACAGAAGTCAG 804

Qy   2043 atactgtgctttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttactc 2102
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    803 ATACTGTGCTTTTTCCAGCAAAGGATCTCAAGGAAAAGGACCTTCATTCAATATTTACTC 744

Qy   2103 atgattctggtctgataacaataaacagttcacaagagcacctaactgttcaggcaaagg 2162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 ATGATTCTGGTCTGATAACAATAAACAGTTCACAAGAGCACCTAACTGTTCAGGCAAAGG 684

Qy   2163 ctccattccatactcctcctgaggaacccaatgaatgtgacttcaagaatatggatagtt 2222
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 CTCCATTCCATACTCCTCCTGAGGAACCCAATGAATGTGACTTCAAGAATATGGATAGTT 624

Qy   2223 taccttctggtaaaatacatcgaaaagtgaaaataatattaggacgaaatagaaaagaaa 2282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 TACCTTCTGGTAAAATACATCGAAAAGTGAAAATAATATTAGGACGAAATAGAAAAGAAA 564

Qy   2283 atctggaaccaaatgctgaatttgataaaagaactgaatttattacacaagaagaaaaca 2342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    563 ATCTGGAACCAAATGCTGAATTTGATAAAAGAACTGAATTTATTACACAAGAAGAAAACA 504

Qy   2343 gaatttgtagttcaccggtacagtctttactagacttgtttcagactagtgaagagaaat 2402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 GAATTTGTAGTTCACCGGTACAGTCTTTACTAGACTTGTTTCAGACTAGTGAAGAGAAAT 444

Qy   2403 cagaatttttgggtttcacaagctacacagaaaagagtggtatatgcaatgttttagata 2462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 CAGAATTTTTGGGTTTCACAAGCTACACAGAAAAGAGTGGTATATGCAATGTTTTAGATA 384

Qy   2463 tttgggaagaggaaaattcagataatctgttaacagcgtttttctcgtccccttcaactt 2522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 TTTGGGAAGAGGAAAATTCAGATAATCTGTTAACAGCGTTTTTCTCGTCCCCTTCAACTT 324

Qy   2523 ctacatttactggctttttagaatttaaaaaaatgcatacttttcagaagtgataaggatca 2582
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 CTACATTTACTGGCTTTTAGAATTTAAAAAATGCATACTTTTCAGAAGTGATAAGGATCA 264

Qy   2583 tattcttgaaatttttataaatatgtatggaaattcttaggatttttttaccagctttgt 2642
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 TATTCTTGAAATTTTTATAAATATGTATGGAAATTCTTAGGATTTTTTTACCAGCTTTGT 204

Qy   2643 ttacagacccaaatgtaaatattaaaaataaatatttgcaattttctacagaattgaata 2702
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 TTACAGACCCAAATGTAAATATTAAAAATAAATATTTGCAATTTTCTACAGAATTGAATA 144

Qy   2703 cctgttaaagaaaaattacagaataaacttgtgactggtcttgttttacattaaaaaaa 2761
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| | | |
Db    143 CCTGTTAAAGAAAAATTACAGAATAAACTTGTGACTGGTCTTGTTTTACATTATATATA 85


RESULT   5
AAI03627/c
ID   AAI03627 standard; DNA; 1915 BP.
XX
AC   AAI03627;
XX
```

```
DT   09-OCT-2001  (first entry)
XX
DE   Probe #3618 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 3618; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and
CC   non-carcinoma tumours.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1915 BP; 608 A; 324 C; 280 G; 703 T; 0 other;

  Query Match             43.0%;  Score 1194.2;  DB 22;  Length 1915;
  Best Local Similarity   99.7%;  Pred. No. 5.2e-258;
  Matches 1196;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1563 aaagaataaaatacagtgttggatccctttctcctgtttctgcaagtgtcctgaaaaaga 1622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1283 AAAGAATAAAATACAGTGTTGGATCCCTTTCTCCTGTTTCTGCAAGTGTCCTGAAAAAGA 1224

Qy   1623 ctgaacaaaaggaaaaagtggaattgcaacatatttctcagaaagattgccaggaagatg 1682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1223 CTGAACAAAAGGAAAAAGTGGAATTGCAACATATTTCTCAGAAAGATTGCCAGGAAGATG 1164

Qy   1683 atacaacagtgaaggagcagaatttcctgtataaagagacccaggaaactgaaaaaaagc 1742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1163 ATACAACAGTGAAGGAGCAGAATTTCCTGTATAAAGAGACCCAGGAAACTGAAAAAAAGC 1104

Qy   1743 tcctgtttatttcagagcccatcccccacccttcaaatgaattgagagggcttaatgaga 1802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1103 TCCTGTTTATTTCAGAGCCCATCCCCCACCCTTCAAATGAATTGAGAGGGCTTAATGAGA 1044

Qy   1803 aaatgagtaataaatgttccatgttaagtacagctgaagatgacataagacagaatttta 1862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   1043 AAATGAGTAATAAATGTTCCATGTTAAGTACAGCTGAAGATGACATAAGACAGAATTTTA 984

Qy   1863 cacagctacctctacataaaaacaaacaggaatgcattcttgacatttccgaacacacat 1922
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    983 CACAGCTACCTCTACATAAAAACAAACAGGAATGCATTCTTGACATTTCCGAACACACAT 924

Qy   1923 taagtgaaaatgacttagaagaactaagggtagatcactataaatgtaacatacaggcat 1982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    923 TAAGTGAAAATGACTTAGAAGAACTAAGGGTAGATCACTATAAATGTAACATACAGGCAT 864

Qy   1983 ctgtacatgtttctgatttcagtacagataatagtggatctcaaccaaaacagaagtcag 2042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    863 CTGTACATGTTTCTGATTTCAGTACAGATAATAGTGGATCTCAACCAAAACAGAAGTCAG 804

Qy   2043 atactgtgctttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttactc 2102
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    803 ATACTGTGCTTTTTCCAGCAAAGGATCTCAAGGAAAAGGACCTTCATTCAATATTTACTC 744

Qy   2103 atgattctggtctgataacaataaacagttcacaagagcacctaactgttcaggcaaagg 2162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    743 ATGATTCTGGTCTGATAACAATAAACAGTTCACAAGAGCACCTAACTGTTCAGGCAAAGG 684

Qy   2163 ctccattccatactcctcctgaggaacccaatgaatgtgacttcaagaatatggatagtt 2222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    683 CTCCATTCCATACTCCTCCTGAGGAACCCAATGAATGTGACTTCAAGAATATGGATAGTT 624

Qy   2223 taccttctggtaaaatacatcgaaaagtgaaaataatattaggacgaaatagaaaagaaa 2282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    623 TACCTTCTGGTAAAATACATCGAAAAGTGAAAATAATATTAGGACGAAATAGAAAAGAAA 564

Qy   2283 atctggaaccaaatgctgaatttgataaaagaactgaatttattacacaagaagaaaaca 2342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    563 ATCTGGAACCAAATGCTGAATTTGATAAAAGAACTGAATTTATTACACAAGAAGAAAACA 504

Qy   2343 gaatttgtagttcaccggtacagtctttactagacttgtttcagactagtgaagagaaat 2402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    503 GAATTTGTAGTTCACCGGTACAGTCTTTACTAGACTTGTTTCAGACTAGTGAAGAGAAAT 444

Qy   2403 cagaatttttgggtttcacaagctacacagaaaagagtggtatatgcaatgttttagata 2462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    443 CAGAATTTTTGGGTTTCACAAGCTACACAGAAAAGAGTGGTATATGCAATGTTTTAGATA 384

Qy   2463 tttgggaagaggaaaattcagataatctgttaacagcgtttttctcgtccccttcaactt 2522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    383 TTTGGGAAGAGGAAAATTCAGATAATCTGTTAACAGCGTTTTTCTCGTCCCCTTCAACTT 324

Qy   2523 ctacatttactggcttttagaatttaaaaaatgcatacttttcagaagtgataaggatca 2582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 CTACATTTACTGGCTTTTAGAATTTAAAAAATGCATACTTTTCAGAAGTGATAAGGATCA 264

Qy   2583 tattcttgaaatttttataaatatgtatggaaattcttaggatttttttaccagctttgt 2642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 TATTCTTGAAATTTTTATAAATATGTATGGAAATTCTTAGGATTTTTTTACCAGCTTTGT 204

Qy   2643 ttacagacccaaatgtaaatattaaaaataaatatttgcaattttctacagaattgaata 2702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 TTACAGACCCAAATGTAAATATTAAAAATAAATATTTGCAATTTTCTACAGAATTGAATA 144

Qy   2703 cctgttaaagaaaaattacagaataaacttgtgactggtcttgttttacattaaaaaaa 2761
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| | | |
Db    143 CCTGTTAAAGAAAAATTACAGAATAAACTTGTGACTGGTCTTGTTTTACATTATATATA 85

RESULT   6
AAH17747
ID   AAH17747 standard; cDNA; 2276 BP.
XX
AC   AAH17747;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:17368.
XX
```

```
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI   Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 17368; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 2276 BP; 793 A; 381 C; 421 G; 681 T; 0 other;


  Query Match            39.6%;  Score 1101.6;  DB 22;  Length 2276;
  Best Local Similarity  99.6%;  Pred. No. 3.2e-237;
  Matches 1104;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy   1654 tatttctcagaaagattgccaggaagatgatacaacagtgaaggagcagaatttcctgta 1713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tatttctcagaaagattgccaggaagatgatacaacagtgaaggagcagaatttcctgta 60

Qy   1714 taaagagacccaggaaactgaaaaaaagctcctgtttatttcagagcccatcccccaccc 1773
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 taaagagacccaggaaactgaaaaaaagctcctgtttatttcagagcccatcccccaccc 120

Qy   1774 ttcaaatgaattgagagggcttaatgagaaaatgagtaataaatgttccatgttaagtac 1833
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ttcaaatgaattgagagggcttaatgagaaaatgagtaataaatgttccatgttaagtac 180

Qy   1834 agctgaagatgacataagacagaattttacacagctacctctacataaaaacaaacagga 1893
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 agctgaagatgacataagacagaattttacacagctacctctacataaaaacaaacagga 240

Qy   1894 atgcattcttgacatttccgaacacacattaagtgaaaatgacttagaagaactaagggt 1953
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 atgcattcttgacatttccgaacacacattaagtgaaaatgacttagaagaactaagggt 300

Qy   1954 agatcactataaatgtaacatacaggcatctgtacatgtttctgatttcagtacagataa 2013
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 agatcactataaatgtaacatacaggcatctgtacatgtttctgatttcagtacagataa 360

Qy   2014 tagtggatctcaaccaaaacagaagtcagatactgtgctttttccagcaaaggatctcaa 2073
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 tagtggatctcaaccaaaacagaagtcagatactgtgctttttccagcaaaggatctcaa 420

Qy   2074 ggaaaaggaccttcattcaatatttactcatgattctggtctgataacaataaacagttc 2133
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ggaaaaggaccttcattcaatatttactcatgattctggtctgataacaataaacagttc 480

Qy   2134 acaagagcacctaactgttcaggcaaaggctccattccatactcctcctgaggaacccaa 2193
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 acaagagcacctaactgttcaggcaaaggctccattccatactcctcctgaggaacccaa 540

Qy   2194 tgaatgtgacttcaagaatatggatagtttaccttctggtaaaatacatcgaaaagtgaa 2253
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    541 tgaatgtgacttcaagaatatggatagtttaccctctggtaaaatacatcgaaaagtgaa 600

Qy   2254 aataatattaggacgaaatagaaaagaaaatctggaaccaaatgctgaatttgataaaag 2313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 aataatattaggacgaaatagaaaagaaaatctggaaccaaatgctgaatttgataaaag 660

Qy   2314 aactgaatttattacacaagaagaaaacagaatttgtagttcaccggtacagtctttact 2373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 aactgaatttattacacaagaagaaaacagaatttgtagttcaccggtacagtctttact 720

Qy   2374 agacttgtttcagactagtgaagagaaatcagaatttttgggtttcacaagctacacaga 2433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 agacttgtttcagactagtgaagagaaatcagaatttttgggtttcacaagctacacaga 780

Qy   2434 aaagagtggtatatgcaatgttttagatatttgggaagaggaaaattcagataatctgtt 2493
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 aaagagtggtatatgcaatgttttagatatttgggaagaggaaaattcagataatctgtt 840

Qy   2494 aacagcgtttttctcgtccccttcaacttctacatttactggcttttagaatttaaaaaa 2553
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 aacagcgtttttctcgtccccttcaacttctacatttactggcttttagaatttaaaaaa 900

Qy   2554 tgcatacttttcagaagtgataaggatcatattcttgaaatttttataaatatgtatgga 2613
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 tgcatacttttcagaagtgataaggatcatattcttgaaatttttataaatatgtatgga 960

Qy   2614 aattcttaggatttttttaccagctttgtttacagacccaaatgtaaatattaaaaataa 2673
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 aattcttaggatttttttaccagctttgtttacagacccaaatgtaaatattaaaaataa 1020

Qy   2674 atatttgcaattttctacagaattgaatacctgttaaagaaaaattacagaataaacttg 2733
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 atatttgcaattttctacagaattgaatacctgttaaagaaaaattacagaataaacttg 1080

Qy   2734 tgactggtcttgttttacattaaaaaaa 2761
           ||||||||||||||||||| | | |
Db   1081 tgactggtcttgttttacattatatata 1108


RESULT   7
AAI22968/c
ID   AAI22968 standard; DNA; 1031 BP.
XX
AC   AAI22968;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #12901 for gene expression analysis in human cervical cell sample.
```

```
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 12901; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1031 BP; 290 A; 181 C; 180 G; 380 T; 0 other;


  Query Match             34.7%;  Score 965;  DB 22;  Length 1031;
  Best Local Similarity   100.0%;  Pred. No. 9.3e-207;
  Matches  965;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1563 aaagaataaaatacagtgttggatccctttctcctgtttctgcaagtgtcctgaaaaaga 1622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    965 AAAGAATAAAATACAGTGTTGGATCCCTTTCTCCTGTTTCTGCAAGTGTCCTGAAAAAGA 906

Qy   1623 ctgaacaaaaggaaaaagtggaattgcaacatatttctcagaaagattgccaggaagatg 1682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    905 CTGAACAAAAGGAAAAAGTGGAATTGCAACATATTTCTCAGAAAGATTGCCAGGAAGATG 846

Qy   1683 atacaacagtgaaggagcagaatttcctgtataaagagacccaggaaactgaaaaaaagc 1742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    845 ATACAACAGTGAAGGAGCAGAATTTCCTGTATAAAGAGACCCAGGAAACTGAAAAAAGC 786

Qy   1743 tcctgtttatttcagagcccatcccccaccttcaaatgaattgagagggcttaatgaga 1802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    785 TCCTGTTTATTTCAGAGCCCATCCCCCACCCTTCAAATGAATTGAGAGGGCTTAATGAGA 726

Qy   1803 aaatgagtaataaatgttccatgttaagtacagctgaagatgacataagacagaatttta 1862
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    725 AAATGAGTAATAAATGTTCCATGTTAAGTACAGCTGAAGATGACATAAGACAGAATTTTA 666

Qy   1863 cacagctacctctacataaaaacaaacaggaatgcattcttgacatttccgaacacacat 1922
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    665 CACAGCTACCTCTACATAAAAACAAACAGGAATGCATTCTTGACATTTCCGAACACACAT 606

Qy   1923 taagtgaaaatgacttagaagaactaagggtagatcactataaatgtaacatacaggcat 1982
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    605 TAAGTGAAAATGACTTAGAAGAACTAAGGGTAGATCACTATAAATGTAACATACAGGCAT 546

Qy   1983 ctgtacatgtttctgatttcagtacagataatagtggatctcaaccaaaacagaagtcag 2042
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    545 CTGTACATGTTTCTGATTTCAGTACAGATAATAGTGGATCTCAACCAAAACAGAAGTCAG 486

Qy   2043 atactgtgctttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttactc 2102
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    485 ATACTGTGCTTTTTCCAGCAAAGGATCTCAAGGAAAAGGACCTTCATTCAATATTTACTC 426

Qy   2103 atgattctggtctgataacaataaacagttcacaagagcacctaactgttcaggcaaagg 2162
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    425 ATGATTCTGGTCTGATAACAATAAACAGTTCACAAGAGCACCTAACTGTTCAGGCAAAGG 366

Qy   2163 ctccattccatactcctcctgaggaacccaatgaatgtgacttcaagaatatggatagtt 2222
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    365 CTCCATTCCATACTCCTCCTGAGGAACCCAATGAATGTGACTTCAAGAATATGGATAGTT 306

Qy   2223 taccttctggtaaaatacatcgaaaagtgaaaataatattaggacgaaatagaaaagaaa 2282
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    305 TACCTTCTGGTAAAATACATCGAAAAGTGAAAATAATATTAGGACGAAATAGAAAAGAAA 246

Qy   2283 atctggaaccaaatgctgaatttgataaaagaactgaatttattacacaagaagaaaaca 2342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    245 ATCTGGAACCAAATGCTGAATTTGATAAAAGAACTGAATTTATTACACAAGAAGAAAACA 186

Qy   2343 gaatttgtagttcaccggtacagtctttactagacttgtttcagactagtgaagagaaat 2402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    185 GAATTTGTAGTTCACCGGTACAGTCTTTACTAGACTTGTTTCAGACTAGTGAAGAGAAAT 126

Qy   2403 cagaatttttgggtttcacaagctacacagaaaagagtggtatatgcaatgttttagata 2462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    125 CAGAATTTTTGGGTTTCACAAGCTACACAGAAAAGAGTGGTATATGCAATGTTTTAGATA 66

Qy   2463 tttgggaagaggaaaattcagataatctgttaacagcgtttttctcgtccccttcaacttt 2522
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65 TTTGGGAAGAGGAAAATTCAGATAATCTGTTAACAGCGTTTTTCTCGTCCCCTTCAACTT 6

Qy   2523 ctaca 2527
           |||||
Db      5 CTACA 1


RESULT   8
AAI08634/c
ID   AAI08634 standard; DNA; 1031 BP.
XX
AC   AAI08634;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #8625 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
```

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -
XX
PS Claim 25; SEQ ID No 8625; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1031 BP; 290 A; 181 C; 180 G; 380 T; 0 other;

Query Match 34.7%; Score 965; DB 22; Length 1031;
Best Local Similarity 100.0%; Pred. No. 9.3e-207;
Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1563 aaagaataaaatacagtgttggatccctttctcctgtttctgcaagtgtcctgaaaaaga 1622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    965 AAAGAATAAAATACAGTGTTGGATCCCTTTCTCCTGTTTCTGCAAGTGTCCTGAAAAAGA 906

Qy   1623 ctgaacaaaaggaaaaagtggaattgcaacatatttctcagaaagattgccaggaagatg 1682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    905 CTGAACAAAAGGAAAAAGTGGAATTGCAACATATTTCTCAGAAAGATTGCCAGGAAGATG 846

Qy   1683 atacaacagtgaaggagcagaatttcctgtataaagagacccaggaaactgaaaaaaagc 1742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    845 ATACAACAGTGAAGGAGCAGAATTTCCTGTATAAAGAGACCCAGGAAACTGAAAAAAAGC 786

Qy   1743 tcctgtttatttcagagcccatcccccacccttcaaatgaattgagagggcttaatgaga 1802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    785 TCCTGTTTATTTCAGAGCCCATCCCCCACCCTTCAAATGAATTGAGAGGGCTTAATGAGA 726

Qy   1803 aaatgagtaataaatgttccatgttaagtacagctgaagatgacataagacagaatttta 1862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    725 AAATGAGTAATAAATGTTCCATGTTAAGTACAGCTGAAGATGACATAAGACAGAATTTTA 666

Qy   1863 cacagctacctctacataaaaacaaacaggaatgcattcttgacatttccgaacacacat 1922
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    665 CACAGCTACCTCTACATAAAAACAAACAGGAATGCATTCTTGACATTTCCGAACACACAT 606

Qy   1923 taagtgaaaatgacttagaagaactaagggtagatcactataaatgtaacatacaggcat 1982
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    605 TAAGTGAAAATGACTTAGAAGAACTAAGGGTAGATCACTATAAATGTAACATACAGGCAT 546

Qy   1983 ctgtacatgtttctgatttcagtacagataatagtggatctcaaccaaaacagaagtcag 2042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    545 CTGTACATGTTTCTGATTTCAGTACAGATAATAGTGGATCTCAACCAAAACAGAAGTCAG 486

Qy   2043 atactgtgctttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttactc 2102
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    485 ATACTGTGCTTTTTCCAGCAAAGGATCTCAAGGAAAAGGACCTTCATTCAATATTTACTC 426

Qy   2103 atgattctggtctgataacaataaacagttcacaagagcacctaactgttcaggcaaagg 2162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    425 ATGATTCTGGTCTGATAACAATAAACAGTTCACAAGAGCACCTAACTGTTCAGGCAAAGG 366

Qy   2163 ctccattccatactcctcctgaggaacccaatgaatgtgacttcaagaatatggatagtt 2222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    365 CTCCATTCCATACTCCTCCTGAGGAACCCAATGAATGTGACTTCAAGAATATGGATAGTT 306

Qy   2223 taccttctggtaaaatacatcgaaaagtgaaaataatattaggacgaaatagaaaagaaa 2282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    305 TACCTTCTGGTAAAATACATCGAAAAGTGAAAATAATATTAGGACGAAATAGAAAAGAAA 246

Qy   2283 atctggaaccaaatgctgaatttgataaaagaactgaatttattacacaagaagaaaaca 2342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    245 ATCTGGAACCAAATGCTGAATTTGATAAAAGAACTGAATTTATTACACAAGAAGAAAACA 186

Qy   2343 gaatttgtagttcaccggtacagtctttactagacttgtttcagactagtgaagagaaat 2402
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    185 GAATTTGTAGTTCACCGGTACAGTCTTTACTAGACTTGTTTCAGACTAGTGAAGAGAAAT 126

Qy   2403 cagaatttttgggtttcacaagctacacagaaaagagtggtatatgcaatgttttagata 2462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    125 CAGAATTTTTGGGTTTCACAAGCTACACAGAAAAGAGTGGTATATGCAATGTTTTAGATA 66

Qy   2463 tttgggaagaggaaaattcagataatctgttaacagcgtttttctcgtccccttcaactt 2522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65 TTTGGGAAGAGGAAAATTCAGATAATCTGTTAACAGCGTTTTTCTCGTCCCCTTCAACTT 6

Qy   2523 ctaca 2527
          |||||
Db      5 CTACA 1
```

RESULT 9
AAH06140
ID AAH06140 standard; cDNA; 903 BP.
XX
AC AAH06140;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:2975.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 1; SEQ ID 2975; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of

```
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 903 BP; 326 A; 165 C; 157 G; 242 T; 13 other;

  Query Match             26.0%;  Score 721.6;  DB 22;  Length 903;
  Best Local Similarity   92.3%;  Pred. No. 3.1e-152;
  Matches  820;  Conservative    0;  Mismatches   56;  Indels   12;  Gaps    6;

Qy   1654 tatttctcagaaagattgccaggaagatgatacaacagt-gaaggagcagaatttcctgt 1712
            ||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db      1 tatttctcagaaagattgccaggaagatgatacaacagtggaaggagcagaatttcctgt 60

Qy   1713 ataaagagacccaggaaactgaaaaaaagctcctgtttatttcagagcccatcccccacc 1772
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ataaagagacccaggaaactgaaaaaaagctcctgtttatttcagagcccatcccccacc 120

Qy   1773 cttcaaatgaattgagagggcttaatgagaaaatgagtaataaatgttccatgttaagta 1832
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 cttcaaatgaattgagagggcttaatgagaaaatgagtaataaatgttccatgttaagta 180

Qy   1833 cagctgaagatgacataagacagaattttacacagctacctctacataaaaacaaacagg 1892
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 cagctgaagatgacataagacagaattttacacagctacctctacataaaaacaaacagg 240

Qy   1893 aatgcattcttgacatttccgaacacacattaagtgaaaatgacttagaagaactaaggg 1952
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    241 aatgcattcttgacatttccgaacacacattaagtgaaaatgacttagaagaactanggg 300

Qy   1953 tagatcactataaatgtaacatacaggcatctgtacatgtttctgatttcagtacagata 2012
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 tagatcactataaatgtaacatacaggcatctgtacatgtttctgatttcagtacagata 360

Qy   2013 atagtggatctcaaccaaaacagaagtcagatactgtgctttttccagcaaaggatctca 2072
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 atagtggatctcaaccaaaacagaagtcagatactgtgctttttccagcaaaggatctca 420

Qy   2073 aggaaaaggaccttcattcaatatttactcatgattctggtctgataacaataaacagtt 2132
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 aggaaaaggaccttcattcaatatttactcatgattctggtctgataacaataaacagtt 480

Qy   2133 cacaagagcacctaactgttcaggcaaaggctccattccatactcctcctgaggaaccca 2192
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 cacaagagcacctaactgttcaggcnaaggctccattccatactcctcctgaggaaccca 540

Qy   2193 atgaatgtgacttcaagaatatggatagtttaccttctggtaaaatacatcgaaaagtga 2252
            ||||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||
Db    541 atgaatgtgacttcaagaatatggatagtttaccctctggtnaaatacatcgaaaagtga 600

Qy   2253 aaataatatta-ggacgaaatagaaaagaaaatctggaaccaaatgctgaatttgataaa 2311
            |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 aaataatattagggacgaaatagaaaagaaaatctggaaccaaatgctgaatttgataaa 660

Qy   2312 agaactgaatttattacacaagaagaaaacagaatttgtagttcaccggtacagtcttta 2371
            |||||||||||||| | | |||||||||| |||||||  ||||||||||||||||||||
Db    661 agaactgaatttattnccccagaagaaaacngaatttgt-tttcaccggtacagtcttta 719

Qy   2372 ctagacttgtttcagactagtgaagagaaatcagaatttttgggtttcacaagctacaca 2431
            ||| |||  ||          ||| | ||||| ||||||||||||||| |||||||| ||
Db    720 ctanactgtttcagactagtggaaaaaaaatcngaatttttgggtttcncaagctaccca 779

Qy   2432 gaaaa--gagtggtatatgcaatgttttagatatttggg-----aagaggaaaattcaga 2484
            |||||   | ||||||||||| |||||||  | || | ||     ||||  ||     |
Db    780 gaaaaagaattggtatatgccatgtttttaanatattggggaaaaagaaaaattccnaat 839

Qy   2485 taatctgttaacagcgttttt--ctcgtccccttcaacttctacattt 2530
            |    ||||||||||||||  |  ||||| | |||||||| |  ||
Db    840 aatctggttaacagcgttttttcccngtcccttccaacttcttccntt 887


RESULT  10
AAF64883
ID   AAF64883 standard; cDNA; 685 BP.
XX
AC   AAF64883;
XX
DT   09-APR-2001  (first entry)
XX
DE   Novel human polynucleotide, SEQ ID NO: 639.
XX
KW   Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW   breast cancer; lung cancer; cancer detection; ss.
XX
OS   Homo sapiens.
XX
PN   WO200102568-A2.
XX
PD   11-JAN-2001.
XX
PF   30-JUN-2000; 2000WO-US18374.
XX
PR   02-JUL-1999;   99US-0142310.
PR   02-JUL-1999;   99US-0142311.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Williams LT,  Escobedo J,  Innis MA,  Garcia PD,  Klinger J,  Kassam A;
PI   Reinhard C,  Randazzo F,  Kennedy GC,  Pot D,  Lamson G,  Drmanac R;
PI   Crkenjakov R,  Drmanac S,  Dickson M,  Labat I,  Leshkowitiz D;
PI   Kita D,  Garcia V,  Jones LW,  Strache-Crain B;
XX
DR   WPI; 2001-091805/10.
XX
PT   Library of polynucleotides for diagnosing a cancerous state of a
PT   mammalian cell and detecting cancer, particularly of the colon or
PT   prostate, comprises 3351 human polynucleotide sequences -
XX
PS   Claim 9; Page 636; 1046pp; English.
XX
CC   The present sequence is one of 3351 sequences in a library of human
CC   polynucleotides. The library is used to detect differentially expressed
CC   genes correlated with a cancerous state of a mammalian cell and can
CC   detect colon, prostate, breast and lung cancer. The library can be used
CC   to produce probes for detection of mRNA and to produce additional copies
CC   of the polynucleotides. The probes can be used for chromosome mapping of
CC   the polynucleotide and for detection of transcription levels. Ribozymes
CC   or antisense oligonucleotides can be generated. The polynucleotides and
CC   their gene products are used as genetic or biochemical markers (e.g. in
CC   blood or tissues) that will detect the earliest changes along the
CC   carcinogenesis pathway and/or monitor the efficacy of therapies and
CC   preventive interventions. The polynucleotides, polypeptides and
CC   antibodies against them can be used in pharmaceutical compositions to
CC   treat the cancers and proliferative disorders such as neoplasia,
CC   dysplasia and hyperplasia.
XX
SQ   Sequence 685 BP; 238 A; 127 C; 125 G; 187 T; 8 other;
```

```
  Query Match            18.5%;  Score 513.8;  DB 22;  Length 685;
  Best Local Similarity  94.9%;  Pred. No. 1e-105;
  Matches  560;  Conservative    0;  Mismatches   24;  Indels    6;  Gaps    3;

Qy     558 tccagggtggaatccaagtcaaaaatgaaaaaaacagaccatctctgaaatctctgaaaa 617
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tccagggtggaatccaagtcaaaaatgaaaaaaacagaccatctctgaaatctctgaaaa 60

Qy     618 ctgataacaggccagaaaaatccaaatgtaagccactttggggaaaagtattttaccttg 677
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ctgataacaggccagaaaaatccaaatgtaagccactttggggaaaagtattttaccttg 120

Qy     678 acttaccttctgtcaccatatctgaaaaacttcaaaaggacattaaggatctgggagggc 737
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 acttaccttctgtcaccatatctgaaaaacttcaaaaggacattaaggatctgggagggc 180

Qy     738 gagttgaagaatttctcagcaaagatatcagttatcttatttcaaataagaaggaagcta 797
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 gagttgaagaatttctcagcaaagatatcagttatcttatttcaaataagaaggaagcta 240

Qy     798 aatttgcacaaaccttgggtcgaatttctcctgtaccaagtccagaatctgcatatactg 857
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 aatttgcacaaaccttgggtcgaatttctcctgtaccaagtccagaatctgcatatactg 300

Qy     858 cagaaaccacttcacctcatcccagccatgatggaagttcatttaagtcaccagacacag 917
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 cagaaaccacttcacctcatcccagccatgatggaagttcatttaagtcaccagacacag 360

Qy     918 tgtgtttaagcagaggaaaattattagttgaaaaagctatcaaggaccatgattttattc 977
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 tgtgtttaagcagaggaaaattattagttgaaaaagctatcaaggaccatgattttattc 420

Qy     978 cttcaaatagtatattatcaaatgccttgtcatggggagtaaaaattcttcatattgatg 1037
           |||||||||||||||||||||||||||||||| ||||| ||||||||||||||||||||
Db     421 cttcaaatagtatattatcaaatgccttgtcatggngagtanaaattcttcatattgatg 480

Qy    1038 acattagatactacattgaacaaaagaaaaaagagttgtatttactcaagaaatcaagta 1097
           |||||||||||||||||||| |||||||||    |   ||||||||| |||    |  |
Db     481 acattagatactacattgaacanaagaaaaagagnt--gatttactcacgaa--tcanga 536

Qy    1098 cttcagtaagagatgggggcaaaagagttggtagtggtgcacaaaaaaca 1147
           ||||| ||  ||||||||   ||| |  | ||||||||| || ||| ||
Db     537 cttcattannagatgggg--caaaaaagttgtagtggtgcccataaacca 584


RESULT  11
AAF66758
ID   AAF66758 standard; cDNA; 422 BP.
XX
AC   AAF66758;
XX
DT   09-APR-2001  (first entry)
XX
DE   Novel human polynucleotide, SEQ ID NO: 2514.
XX
KW   Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW   breast cancer; lung cancer; cancer detection; ss.
XX
OS   Homo sapiens.
XX
PN   WO200102568-A2.
XX
PD   11-JAN-2001.
XX
PF   30-JUN-2000; 2000WO-US18374.
XX
PR   02-JUL-1999;   99US-0142310.
PR   02-JUL-1999;   99US-0142311.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Williams LT,  Escobedo J,  Innis MA,  Garcia PD,  Klinger J,  Kassam A;
PI   Reinhard C,  Randazzo F,  Kennedy GC,  Pot D,  Lamson G,  Drmanac R;
PI   Crkenjakov R,  Drmanac S,  Dickson M,  Labat I,  Leshkowitiz D;
PI   Kita D,  Garcia V,  Jones LW,  Strache-Crain B;
XX
DR   WPI; 2001-091805/10.
XX
PT   Library of polynucleotides for diagnosing a cancerous state of a
PT   mammalian cell and detecting cancer, particularly of the colon or
PT   prostate, comprises 3351 human polynucleotide sequences -
XX
PS   Claim 9; Page 915; 1046pp; English.
XX
CC   The present sequence is one of 3351 sequences in a library of human
CC   polynucleotides. The library is used to detect differentially expressed
CC   genes correlated with a cancerous state of a mammalian cell and can
CC   detect colon, prostate, breast and lung cancer. The library can be used
CC   to produce probes for detection of mRNA and to produce additional copies
CC   of the polynucleotides. The probes can be used for chromosome mapping of
CC   the polynucleotide and for detection of transcription levels. Ribozymes
CC   or antisense oligonucleotides can be generated. The polynucleotides and
CC   their gene products are used as genetic or biochemical markers (e.g. in
CC   blood or tissues) that will detect the earliest changes along the
CC   carcinogenesis pathway and/or monitor the efficacy of therapies and
CC   preventive interventions. The polynucleotides, polypeptides and
CC   antibodies against them can be used in pharmaceutical compositions to
CC   treat the cancers and proliferative disorders such as neoplasia,
CC   dysplasia and hyperplasia.
XX
SQ   Sequence 422 BP; 147 A; 77 C; 75 G; 123 T; 0 other;


  Query Match            14.8%;  Score 411;  DB 22;  Length 422;
  Best Local Similarity  100.0%;  Pred. No. 9.2e-83;
  Matches  411;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     663 aagtattttaccttgacttaccttctgtcaccatatctgaaaaacttcaaaaggacatta 722
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12 aagtattttaccttgacttaccttctgtcaccatatctgaaaaacttcaaaaggacatta 71

Qy     723 aggatctgggagggcgagttgaagaatttctcagcaaagatatcagttatcttatttcaa 782
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72 aggatctgggagggcgagttgaagaatttctcagcaaagatatcagttatcttatttcaa 131

Qy     783 ataagaaggaagctaaatttgcacaaaccttgggtcgaatttctcctgtaccaagtccag 842
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     132 ataagaaggaagctaaatttgcacaaaccttgggtcgaatttctcctgtaccaagtccag 191

Qy     843 aatctgcatatactgcagaaaccacttcacctcatcccagccatgatggaagttcattta 902
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     192 aatctgcatatactgcagaaaccacttcacctcatcccagccatgatggaagttcattta 251

Qy     903 agtcaccagacacagtgtgtttaagcagaggaaaattattagttgaaaaagctatcaagg 962
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     252 agtcaccagacacagtgtgtttaagcagaggaaaattattagttgaaaaagctatcaagg 311

Qy     963 accatgattttattccttcaaatagtatattatcaaatgccttgtcatggggagtaaaaa 1022
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     312 accatgattttattccttcaaatagtatattatcaaatgccttgtcatggggagtaaaaa 371

Qy    1023 ttcttcatattgatgacattagatactacattgaacaaaagaaaaaagagt 1073
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     372 ttcttcatattgatgacattagatactacattgaacaaaagaaaaaagagt 422


RESULT  12
AAF64912
ID   AAF64912 standard; cDNA; 403 BP.
XX
AC   AAF64912;
XX
DT   09-APR-2001  (first entry)
XX
DE   Novel human polynucleotide, SEQ ID NO: 668.
```

XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitiz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 640; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 403 BP; 144 A; 82 C; 76 G; 101 T; 0 other;

Query Match 13.8%; Score 384.4; DB 22; Length 403;
Best Local Similarity 99.7%; Pred. No. 8.2e-77;
Matches 385; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy    555 atttccagggtggaatccaagtcaaaaatgaaaaaaacagaccatctctgaaatctctga 614
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     18 agttccagggtggaatccaagtcaaaaatgaaaaaaacagaccatctctgaaatctctga 77

Qy    615 aaactgataacaggccagaaaaatccaaatgtaagccactttggggaaaagtattttacc 674
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     78 aaactgataacaggccagaaaaatccaaatgtaagccactttggggaaaagtattttacc 137

Qy    675 ttgacttaccttctgtcaccatatctgaaaaacttcaaaaggacattaaggatctgggag 734
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    138 ttgacttaccttctgtcaccatatctgaaaaacttcaaaaggacattaaggatctgggag 197

Qy    735 ggcgagttgaagaatttctcagcaaagatatcagttatcttatttcaaataagaaggaag 794
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 ggcgagttgaagaatttctcagcaaagatatcagttatcttatttcaaataagaaggaag 257

Qy    795 ctaaatttgcacaaaccttgggtcgaatttctcctgtaccaagtccagaatctgcatata 854
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 ctaaatttgcacaaaccttgggtcgaatttctcctgtaccaagtccagaatctgcatata 317

Qy    855 ctgcagaaaccacttcacctcatcccagccatgatggaagttcatttaagtcaccagaca 914
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 ctgcagaaaccacttcacctcatcccagccatgatggaagttcatttaagtcaccagaca 377

Qy    915 cagtgtgtttaagcagaggaaaatta 940
          |||||||||||||||||||||||||
Db    378 cagtgtgtttaagcagaggaaaatta 403
```

RESULT 13
AAA43876
ID AAA43876 standard; cDNA; 318 BP.
XX
AC AAA43876;
XX
DT 21-AUG-2000 (first entry)
XX
DE Human secreted expressed sequence tag SEQ ID NO:451.
XX
KW Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;
KW expressed sequence tag; EST; probe; chemotactic; proliferative;
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KW antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;
KW antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW central nervous system disorder; Alzheimer's disease; stroke;
KW Parkinson's disease; Huntington's disease; coagulation disorder;
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW tumour; infection; depression; psoriasis; ss.
XX
OS Homo sapiens.
XX
PN WO200021991-A1.
XX
PD 20-APR-2000.
XX
PF 15-OCT-1999; 99WO-US24206.
XX
PR 15-OCT-1998; 98US-0104436.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR;
XX
DR WPI; 2000-317938/27.
XX
PT Isolated polynucleotides, and encoded proteins, comprising secreted
PT expressed sequence tags (sESTs), useful for treating various disorders
PT such as autoimmune, infectious, and central nervous system disorders -
XX
PS Claim 1; Page 321; 803pp; English.
XX
CC AAA43426 to AAA45925 represent specifically claimed secreted expressed
CC sequence tags (sESTs), isolated from human, mouse, chicken and rat
CC tissue sources. The sESTs can have a range of activities depending on
CC the tissues they were isolated from. The activities include:
CC chemotactic; proliferative; immunomodulatory; haematopoietic;
CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
CC antiasthmatic; vulnerary; antiulcer; osteopathic; neuroprotective;
CC nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;
CC anticonvulsant; and antidepressant. The sESTs can be used for gene
CC therapy and in vaccines. The sESTs are useful as probes for the
CC identification and isolation of full-length cDNAs and genomic DNA
CC molecules which correspond to the sESTs. Proteins encoded by the sESTs
CC are useful in assays for determining biological activity and raising
CC antibodies. They may be useful for treatment of autoimmune disorders

```
CC   (multiple sclerosis, insulin dependent diabetes), allergic conditions
CC   (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,
CC   osteoporosis, osteoarthritis, central nervous system disorders
CC   (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
CC   disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
CC   disease), tumours, bacterial, fungal or viral infections, depression and
CC   psoriasis. AAA45926 to AAA45931 represent linker variants which are given
CC   in the exemplification of the present invention.
XX
SQ   Sequence 318 BP; 108 A; 69 C; 54 G; 86 T; 1 other;


  Query Match             10.6%;  Score 294.6;  DB 21;  Length 318;
  Best Local Similarity   96.8%;  Pred. No. 9.9e-57;
  Matches  300;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1873 tctacataaaaacaaacaggaatgcattcttgacatttccgaacacacattaagtgaaaa 1932
           ||  | |  |  |  ||||||||||||||||||||||||||||||||||||||||||||||||
Db      5 tcggccttcatgcctacaggaatgcattcttgacatttccgaacacacattaagtgaaaa 64

Qy   1933 tgacttagaagaactaagggtagatcactataaatgtaacatacaggcatctgtacatgt 1992
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65 tgacttagaagaactaagggtagatcactataaatgtaacatacaggcatctgtacatgt 124

Qy   1993 ttctgatttcagtacagataatagtggatctcaaccaaaacagaagtcagatactgtgct 2052
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    125 ttctgatttcagtacanataatagtggatctcaaccaaaacagaagtcagatactgtgct 184

Qy   2053 ttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttactcatgattctgg 2112
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    185 ttttccagcaaaggatctcaaggaaaaggaccttcattcaatatttactcatgattctgg 244

Qy   2113 tctgataacaataaacagttcacaagagcacctaactgttcaggcaaaggctccattcca 2172
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    245 tctgataacaataaacagttcacaagagcacctaactgttcaggcaaaggctccattcca 304

Qy   2173 tactcctcct 2182
          ||||||||||
Db    305 tactcctcct 314


RESULT  14
AAF64832
ID   AAF64832 standard; cDNA; 397 BP.
XX
AC   AAF64832;
XX
DT   09-APR-2001  (first entry)
XX
DE   Novel human polynucleotide, SEQ ID NO: 588.
XX
KW   Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW   breast cancer; lung cancer; cancer detection; ss.
XX
OS   Homo sapiens.
XX
PN   WO200102568-A2.
XX
PD   11-JAN-2001.
XX
PF   30-JUN-2000; 2000WO-US18374.
XX
PR   02-JUL-1999;   99US-0142310.
PR   02-JUL-1999;   99US-0142311.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI   Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI   Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitiz D;
PI   Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR   WPI; 2001-091805/10.
XX
PT   Library of polynucleotides for diagnosing a cancerous state of a
PT   mammalian cell and detecting cancer, particularly of the colon or
PT   prostate, comprises 3351 human polynucleotide sequences -
XX
PS   Claim 9; Page 629; 1046pp; English.
XX
CC   The present sequence is one of 3351 sequences in a library of human
CC   polynucleotides. The library is used to detect differentially expressed
CC   genes correlated with a cancerous state of a mammalian cell and can
CC   detect colon, prostate, breast and lung cancer. The library can be used
CC   to produce probes for detection of mRNA and to produce additional copies
CC   of the polynucleotides. The probes can be used for chromosome mapping of
CC   the polynucleotide and for detection of transcription levels. Ribozymes
CC   or antisense oligonucleotides can be generated. The polynucleotides and
CC   their gene products are used as genetic or biochemical markers (e.g. in
CC   blood or tissues) that will detect the earliest changes along the
CC   carcinogenesis pathway and/or monitor the efficacy of therapies and
CC   preventive interventions. The polynucleotides, polypeptides and
CC   antibodies against them can be used in pharmaceutical compositions to
CC   treat the cancers and proliferative disorders such as neoplasia,
CC   dysplasia and hyperplasia.
XX
SQ   Sequence 397 BP; 123 A; 48 C; 70 G; 156 T; 0 other;


  Query Match             7.6%;  Score 211.4;  DB 22;  Length 397;
  Best Local Similarity   79.2%;  Pred. No. 4.8e-38;
  Matches  289;  Conservative    0;  Mismatches   71;  Indels    5;  Gaps    3;

Qy    951 aagctatcaaggaccatgattttattc-cttcaaatagtatattatcaaatgccttgtca 1009
          | |  |||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      4 acgagatcaaggaccatgattttattctcttcaaatagtatattatcaaatgccttgtca 63

Qy   1010 tggggagtaaaaattcttcatattgatgacattagatactacattgaacaaaagaaaaaa 1069
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 tggggagtaaaaattcttcatattgatgacattagatactacattgaacaaaagaaaaaa 123

Qy   1070 gagttgtatttactcaagaaatcaagtacttcagtaagagatgggggcaaaagagttggt 1129
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    124 gagttgtatttactcaagaaatcaagtacttcagtaagagatgggggcaaaagagttggt 183

Qy   1130 agtggtgcacaaaaaacaagaacaggaagactcaaaaagccttttgtaaaggtggaagat 1189
          |||||||||||  ||  |  |||   ||||| | ||| ||||||||  | ||| |||||||
Db    184 agtggtgcacacaattcttgaa---gaagatttaaatagccttttttgatagggggaagat 240

Qy   1190 atgagccaactttataggccattttatcttcagctgaccaatatgccttttataaattat 1249
          ||| |||| |||||| |  |  ||| | |||       |  || |    || || |||||
Db    241 atgtgccatctttattgtgccattt-ttttctttatgtctttaaggtggttttatattat 299

Qy   1250 tctattcagaagccctgcagtccatttgatgtagacaagccatctagtatgcaaaagcaa 1309
          ||| |  |||| |||   |    ||||     | | |    || |  ||||  |||
Db    300 tctttgtagaatcccactatggtatttttataatatattgtatttttttatgggaaatttt 359

Qy   1310 actca 1314
          ||||
Db    360 tctca 364


RESULT  15
AAI16129
ID   AAI16129 standard; DNA; 468 BP.
XX
AC   AAI16129;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #6062 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
```

OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID No 6062; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 468 BP; 146 A; 92 C; 92 G; 138 T; 0 other;

Query Match 6.6%; Score 184.2; DB 22; Length 468;
Best Local Similarity 98.4%; Pred. No. 6.5e-32;
Matches 186; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy    736 gcgagttgaagaatttctcagcaaagatatcagttatcttatttcaaataagaaggaagc 795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    274 gcgagttgaagaatttctcagcaaagatatcagttatcttatttcaaataagaaggaagc 333

Qy    796 taaatttgcacaaaccttgggtcgaatttctcctgtaccaagtccagaatctgcatatac 855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 taaatttgcacaaaccttgggtcgaatttctcctgtaccaagtccagaatctgcatatac 393

Qy    856 tgcagaaaccacttcacctcatcccagccatgatggaagttcatttaagtcaccagacac 915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 tgcagaaaccacttcacctcatcccagccatgatggaagttcatttaagtcaccagacac 453

Qy    916 agtgtgttt 924
          |||  || |
Db    454 agtaagtct 462
```

Search completed: December 27, 2001, 20:29:47
Job time: 12842 sec